

Wintersemester 2006/07

# **Fundamente der Computational Intelligence**

**(Vorlesung)**

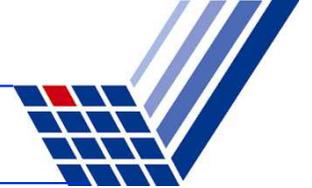
**Dozent: Prof. Dr. Günter Rudolph**

**Vertretung: Boris Naujoks**

Fachbereich Informatik

Lehrstuhl für Algorithm Engineering (LS XI)

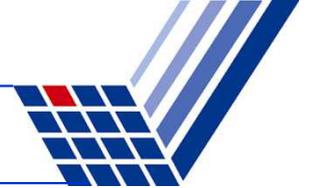




# Multikriterielle Optimierung mit evolutionären Algorithmen

## Übersicht:

1. Beispiel
2. Historie
3. Grundlagen
4. Einfache, aggregierende Ansätze
5. Standardverfahren NSGA-II,  $\epsilon$ -MOEA
6. Vergleich von Ergebnissen: Metriken
7. SMS-EMOA
8. Ausblick, offene Fragen etc.



### Beispiel: Autokauf (gebraucht)

	VW	Opel	Ford	Toyota	
Preis (T€)	16	14	15	13	→ Min!
Verbrauch (l/100km)	7,2	7,0	7,5	7,8	→ Min!
Leistung (kW)	65	55	58	55	→ Max!

3 Ziele, 4 Alternativen → beste Alternative?



# Historische Entwicklung zu EMOA (Evolutionary Multi-Objective Algorithms)

1848 – 1923: Vilfredo Federico Pareto,  
italienischer Ingenieur, Ökonom und Soziologe

1985: J. D. Schaffer, VEGA (1. ICGA)

1990: F. Kursawe, multikriterielle diploide ES (1. PPSN)

1995: C. Fonseca u. P. Fleming: Übersichtsartikel in ECJ (MIT)

2001: International Conference on EMO (Evol- Multi.-obj. Opt.), Zürich  
(Springer-Verlag), seitdem alle 2 Jahre



### Historische Entwicklung in Dortmund

1990: F. Kursawe, multikriterielle diploide ES (1. PPSN)

1998: Räuber-Beute-System von Laumanns (Rudolph und Schwefel)

2000: Theoretische Ergebnisse von Rudolph (und Agapie)

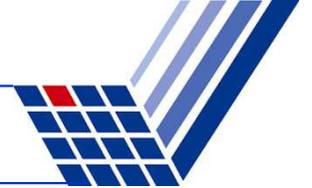
2005: SMS-EMOA von Beume, Emmerich und Naujoks

### Weitere (wichtige) Algorithmen:

2000: Non-dominated Sorting GA (2nd ed., NSGA-II, Deb et al.)

2001: Strength-Pareto EA (2nd ed., SPEA2, Zitzer et al.)

2003:  $\epsilon$ -MOEA (Deb et al., Laumanns et al.)



### Multikriterielle Optimierung

$$f : X \subset \mathbb{R}^n \rightarrow S \subset \mathbb{R}^d$$
$$x, y \in X, \quad a := f(x) \in S, \quad b := f(y) \in S$$

- Konzept der *Pareto Dominanz*:  $x$  dominiert  $y$  (Minimierung)

$$x \prec_p y \Leftrightarrow \forall i : f_i(x) \leq f_i(y) \wedge$$
$$\exists j : f_j(x) < f_j(y)$$

- Menge nicht-dominierter Lösungen: Pareto-Menge, *Pareto-Front*

$$\{x \in X / \nexists y \in X : y \prec_p x\} \quad \{a = f(x) \in S / \nexists y \in X : y \prec_p x\}$$



Optimierung bzgl. mehrerer Kriterien/Ziele

Bsp. Laptop: Gewicht, Preis, Prozessorleistung optimieren

Zielfunktionen  $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_d(\mathbf{x})) \rightarrow \min!$       $\mathbf{f}: X \rightarrow S \in \mathbb{R}^d$

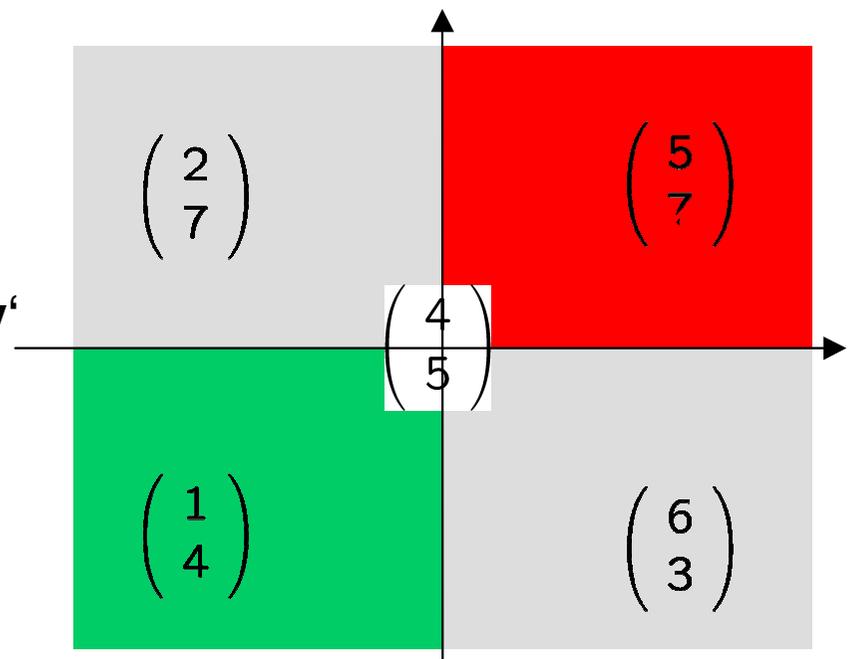
nicht 1 Optimum, sondern Menge optimaler Kompromisslösungen

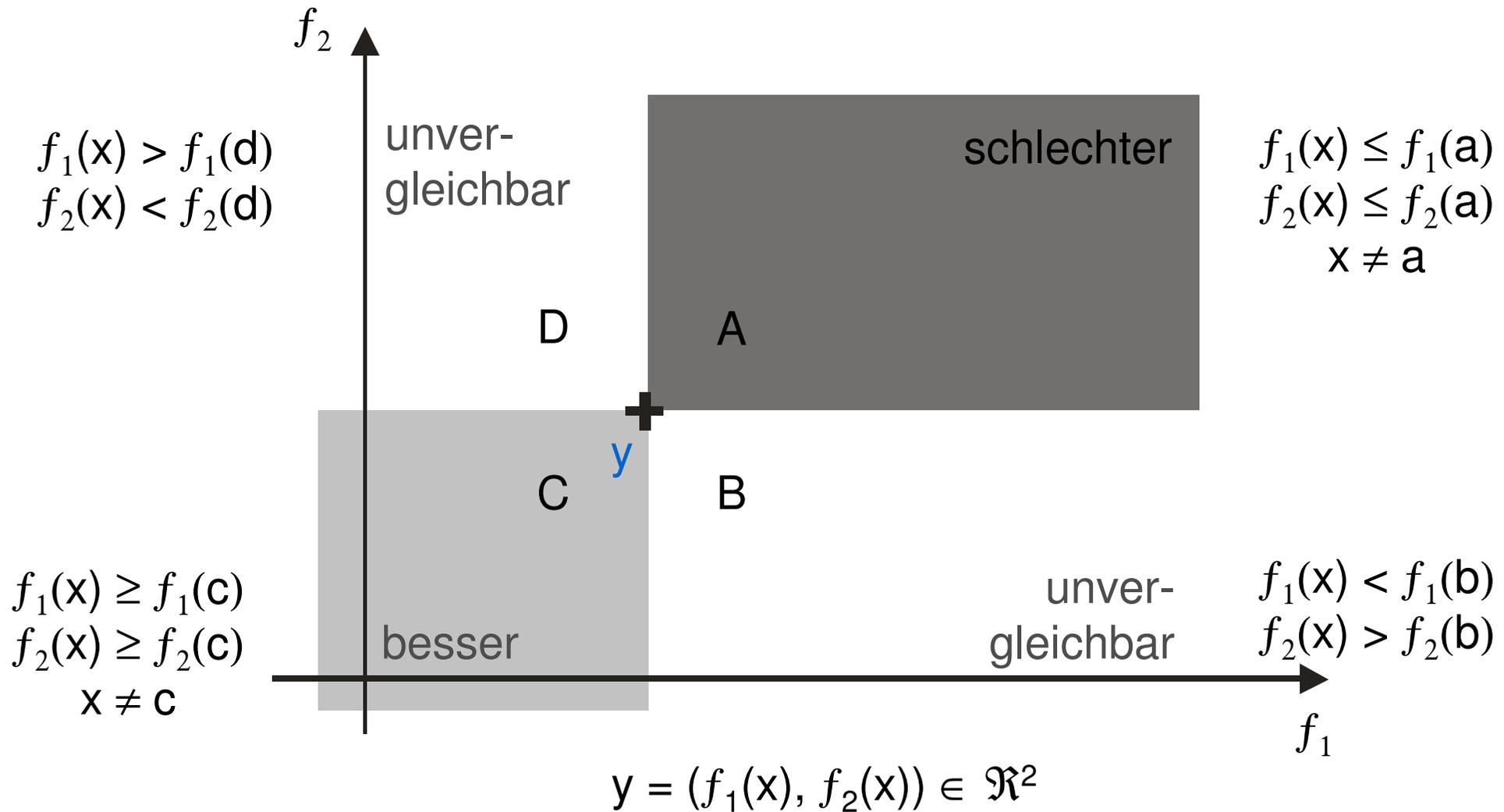
**Pareto-Dominanz:**

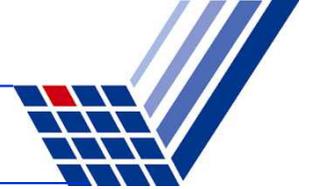
partielle Ordnung zwischen Zielvektoren

$$\mathbf{y} \prec_p \mathbf{y}' \Leftrightarrow \forall i \in \{1, \dots, d\} : y_i \leq y'_i \text{ und } \mathbf{y} \neq \mathbf{y}'$$

unvergl. mit Anteil von  $1/2^{d-1}$







Nicht-dominierte Menge  $ND(A)$ : minimale Elemente in Menge  $A$   
 $\prec_p$

Zielsetzung der Mehrzieloptimierung

Annäherung an Pareto-Front, Abdeckung von Pareto-Front

wiederum multikriteriell

Optimierung durch evolutionäre Algorithmen (EA)

Ergebnis von EA: Approximation der Pareto-Front

Menge von Lösungen

Anwender wählt *nach Optimierung 1* zu realisierende Lösung

$PF_{true}$  - Wahre Pareto - Front

$PF_{known}$  - Bekannte Pareto - Front

$PF_{cur}$  - Momentane Pareto - Front (während Optimierung)



### Aggregierende Ansätze

- Klassische Herangehensweise
- Unterschiedliche Kriterien zu einem aggregieren
- EAs können Konzept der *Pareto Dominanz* direkt umsetzen
- Menge mehrerer Alternativlösungen (Population)

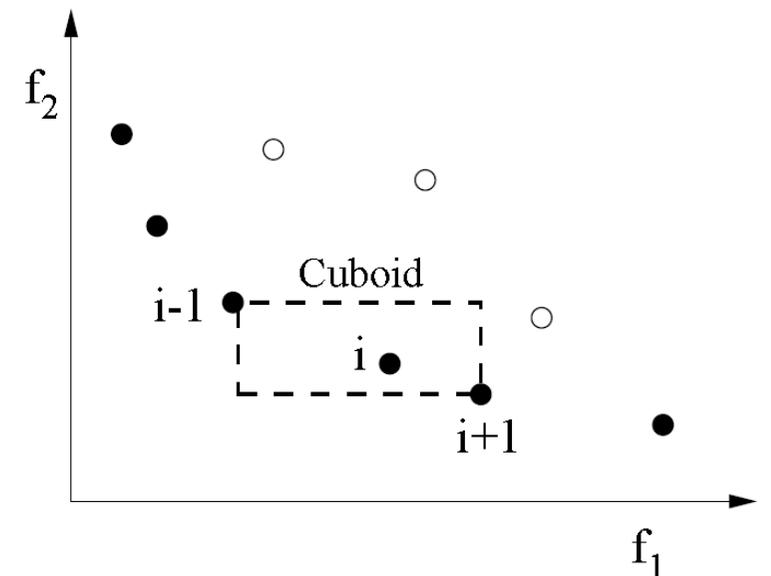
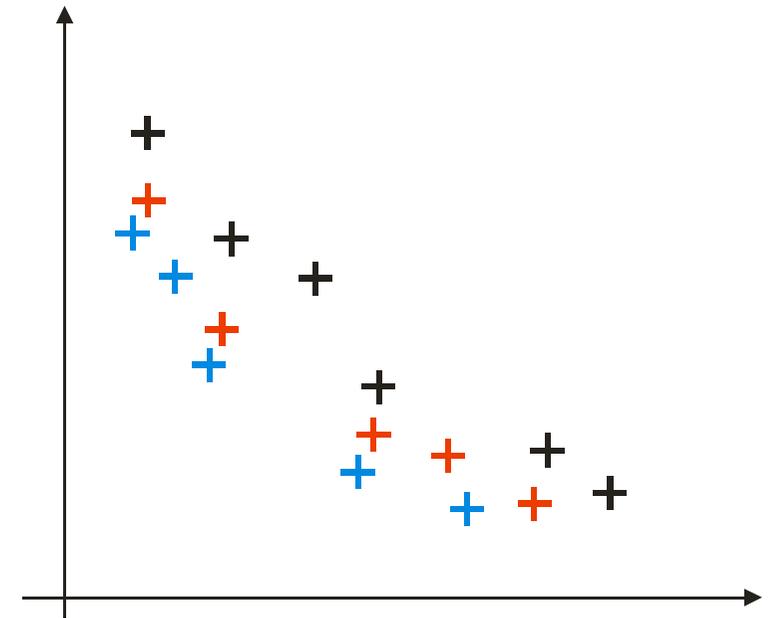
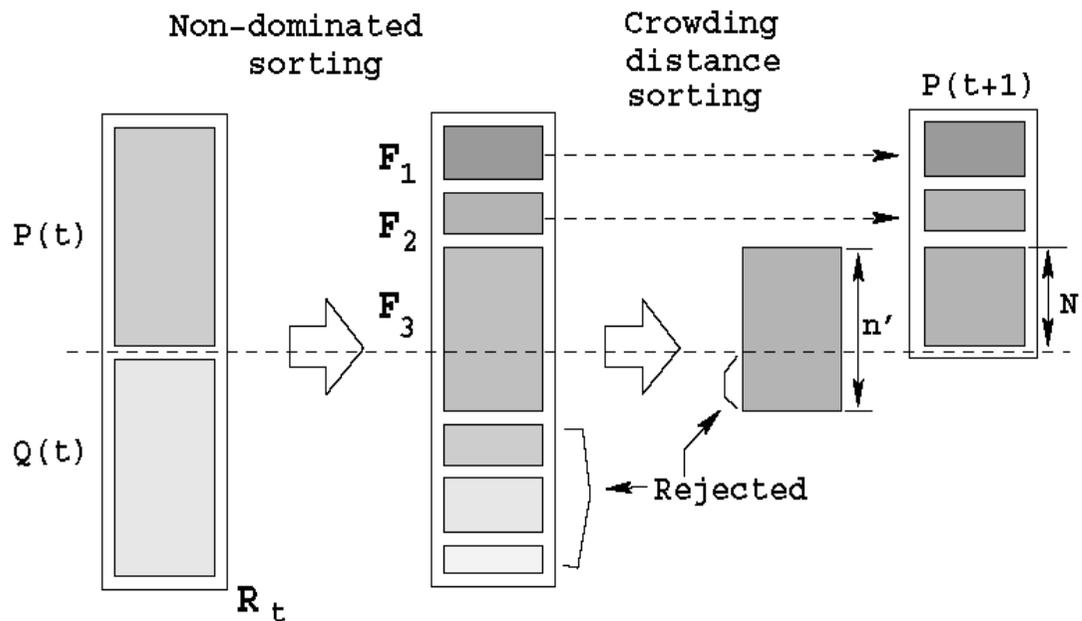
$$f(x) = \sum_i w_i f_i(x)$$

$$f(x) = \min(\max_i f_i(x))$$



## NSGA-II

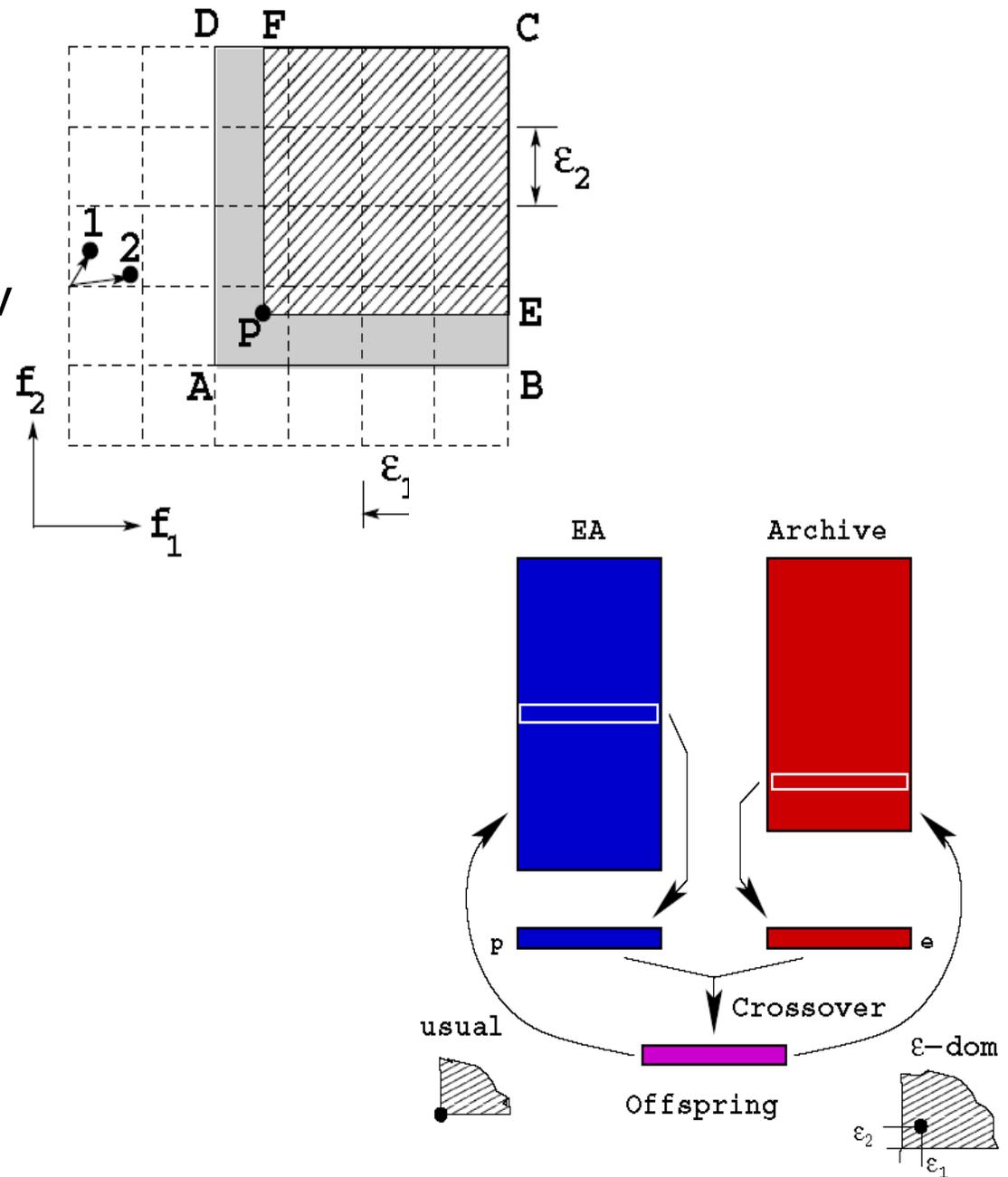
- GA-basiert (Kodierung, Mutation, etc.)
- „Crowding“ Distanz zur Variationserhaltung
- Nicht-dominiertes Sortieren





## $\epsilon$ -MOEA

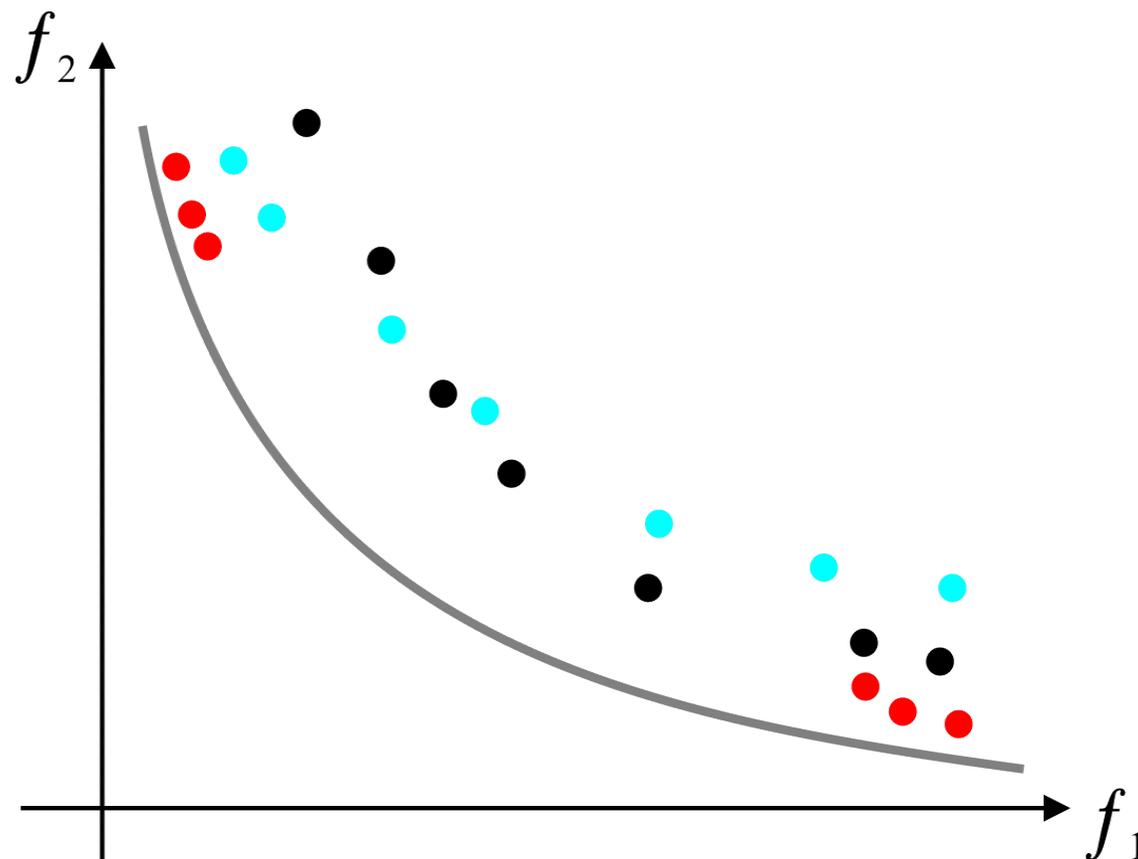
- 2 Individuenmengen im Einsatz: Population und Archiv
- $\epsilon$ -Dominanz ergänzt Prinzip der Pareto-Dominanz
- $\epsilon$ -Dominanz nur im Archiv im Einsatz
- „Steady-state“ Ansatz
- Elitismus: Gute Lösungen gehen nicht verloren





### Welcher EMOA löst mein Problem besser?

- Ergebnismengen verschiedener EA vergleichen



Mengen A, B qualitativ  
unvergleichbar  $\Leftrightarrow$   
Element in A ohne  
Vergleichspartner in B

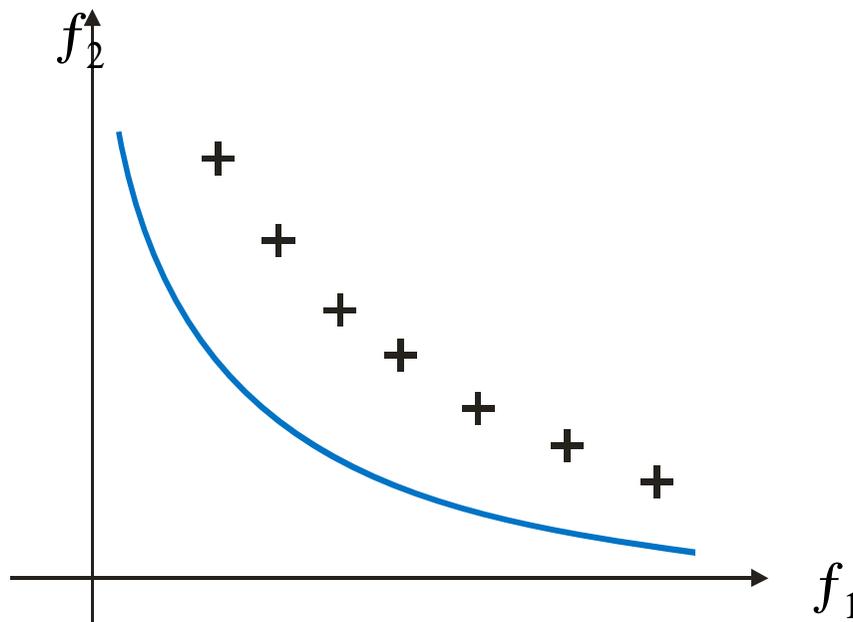
Hilfsmittel: quantitative  
Qualitätsmaße

Bewertung bzgl. beliebiger  
Eigenschaften



### Güte von Mengen

- Konvergenz: Abstand zur Pareto-Front  
Diversität: Verteilung über gesamte Front
- trivial: A besser als B, falls alle Punkte in B dominiert von A
- Bewertung durch Qualitätsmaße: Konvergenz-Metrik, S-Metrik



— Pareto-Front

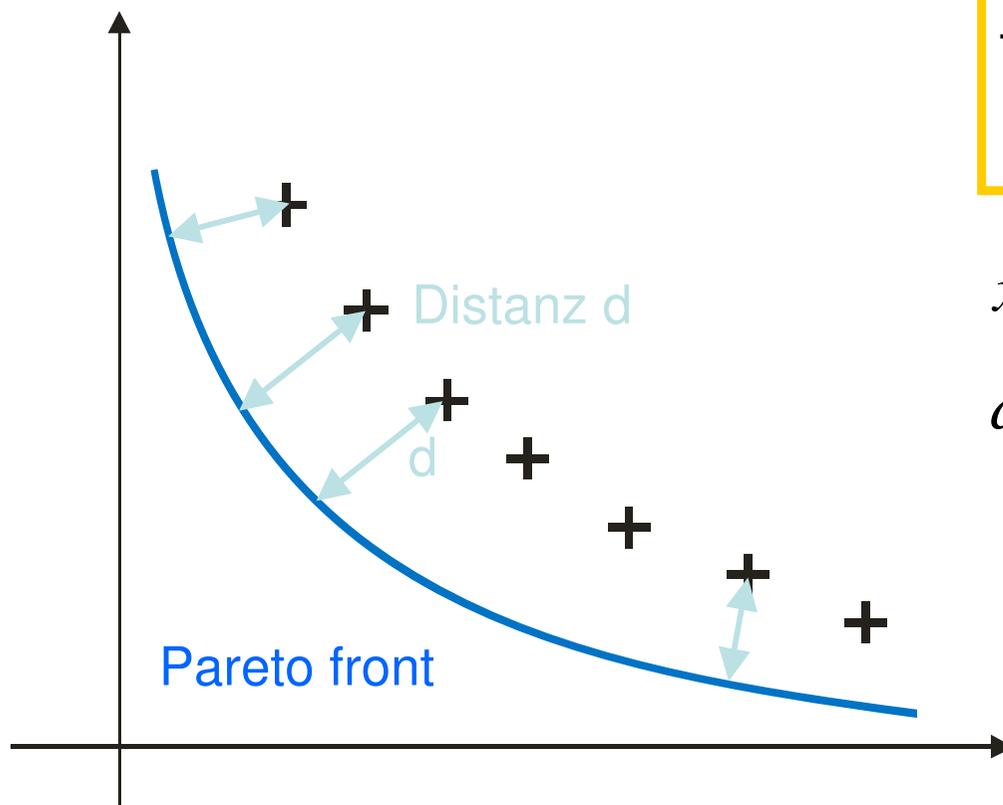
+ Approximierte Menge:  
gefundene,  
nicht-dominierte  
Lösungen

In Spezialvorlesung:  
*Outperformance-Relationen*



## Konvergenz Maß

- Mittelwert der Distanzen der nicht-dominierten Lösungen zur Pareto-Front



$$\frac{1}{|PF_{cur}|} \cdot \sum_{j=1}^{|PF_{cur}|} d(x_i, y_j)$$

$$x_i \in PF_{known}, y_j \in PF_{cur}$$

$$d(x_i, y_j) = \min_{k=1}^{|PF_{known}|} d(x_k, y_j)$$

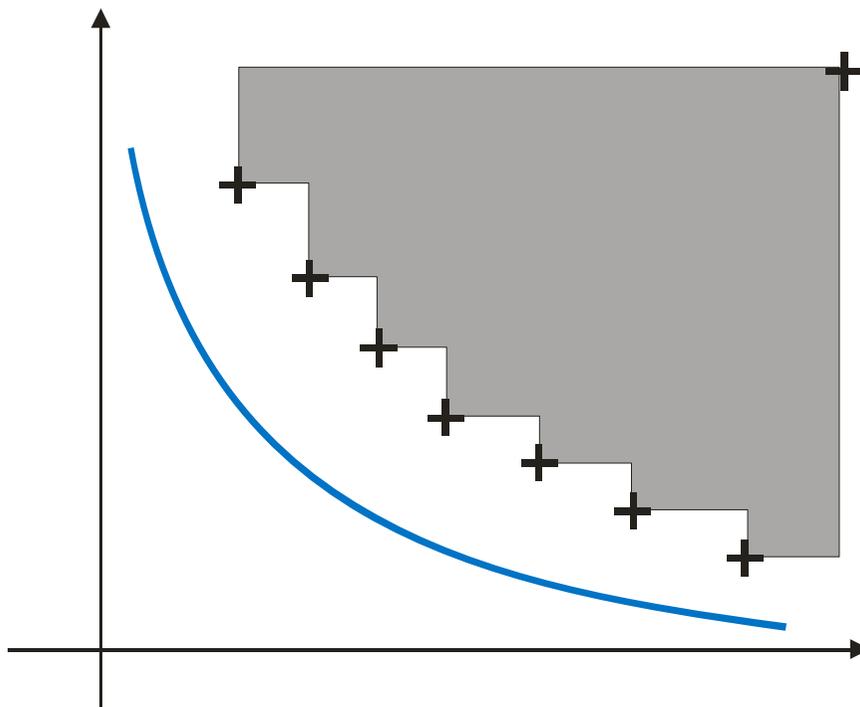
**Aber Vorsicht:**  
Gibt's andersrum ebenfalls



## S-Metrik, Hypervolumen

Zitzler, Thiele, ETH Zürich 1998

- dominiertes Hypervolumen bis Referenzpunkt
- belohnt Annäherung zur und Abdeckung von Pareto-Front
- Hypervolumen des dominierten Raums bzgl. Referenzpunkt  $y_{ref}$
- unäres Maß  $S(M)$
- Definition über Lebesgue-Maß  $\Lambda$  (Vereinigung von Hyperkuben)



$y_{ref}$

$$S(M) := \Lambda( \{ \cup_i c^{(i)} \mid y^{(i)} \in M \} )$$

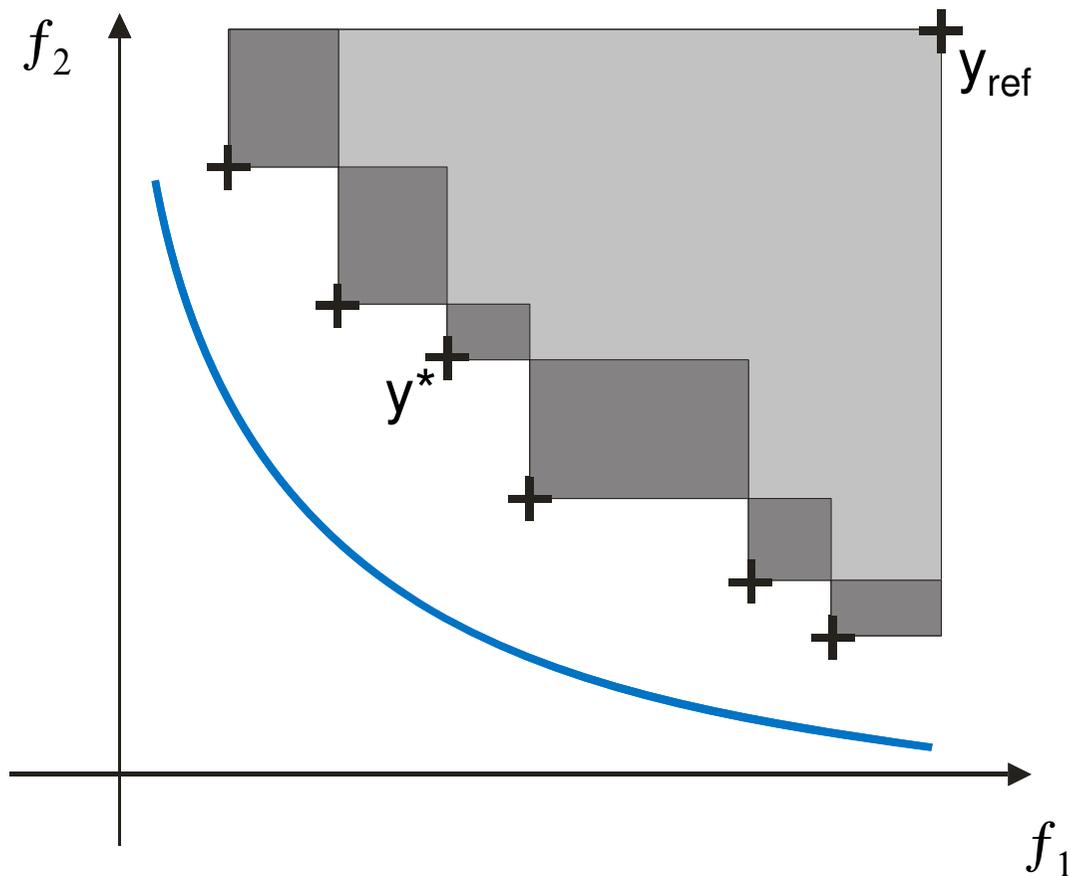
$c^{(i)}$ : Vereinigung von Kuben  
dominiert von  $y^{(i)}$ ,  
dominieren  $y_{ref}$

Exponentielle Komplexität



## S-Metrik-Beitrag

Welche Lösungen sind die wertvollsten?



S-Metrik-Beitrag  $\Delta_S(y, M)$   
von  $y$  zu  $M$

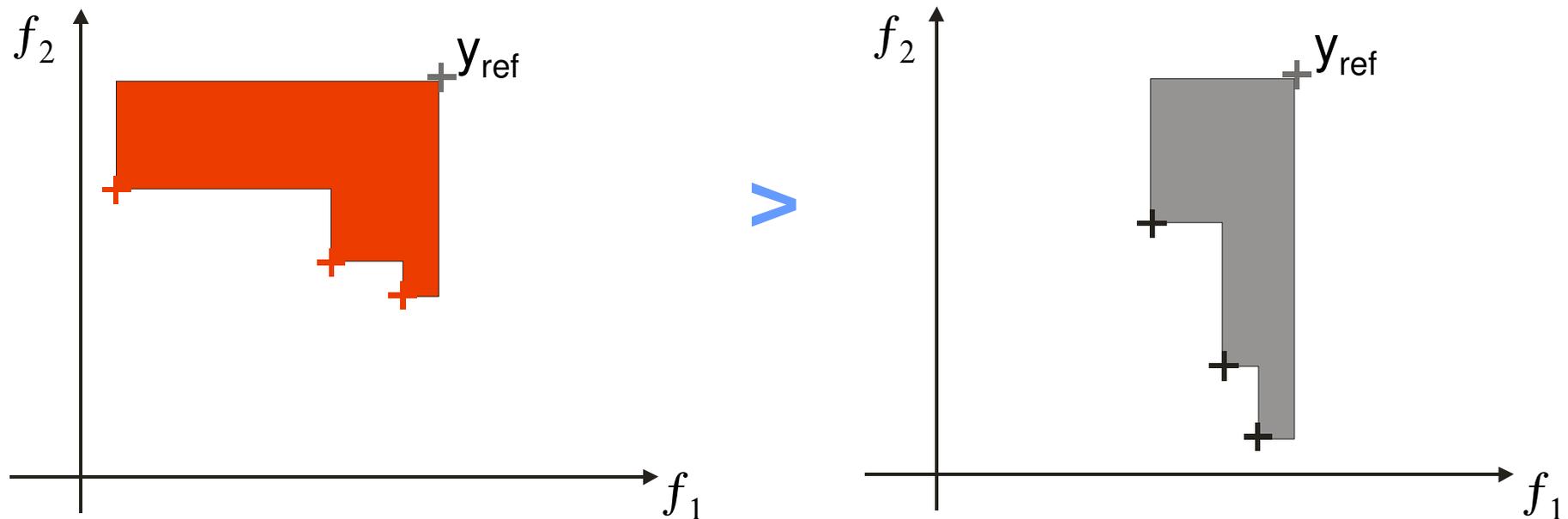
Hypervolumen, in  $M$   
exklusiv dominiert von  $y$

*hier:*  
 $y^*$  kleinster Beitrag



## Eigenschaften der S-Metrik

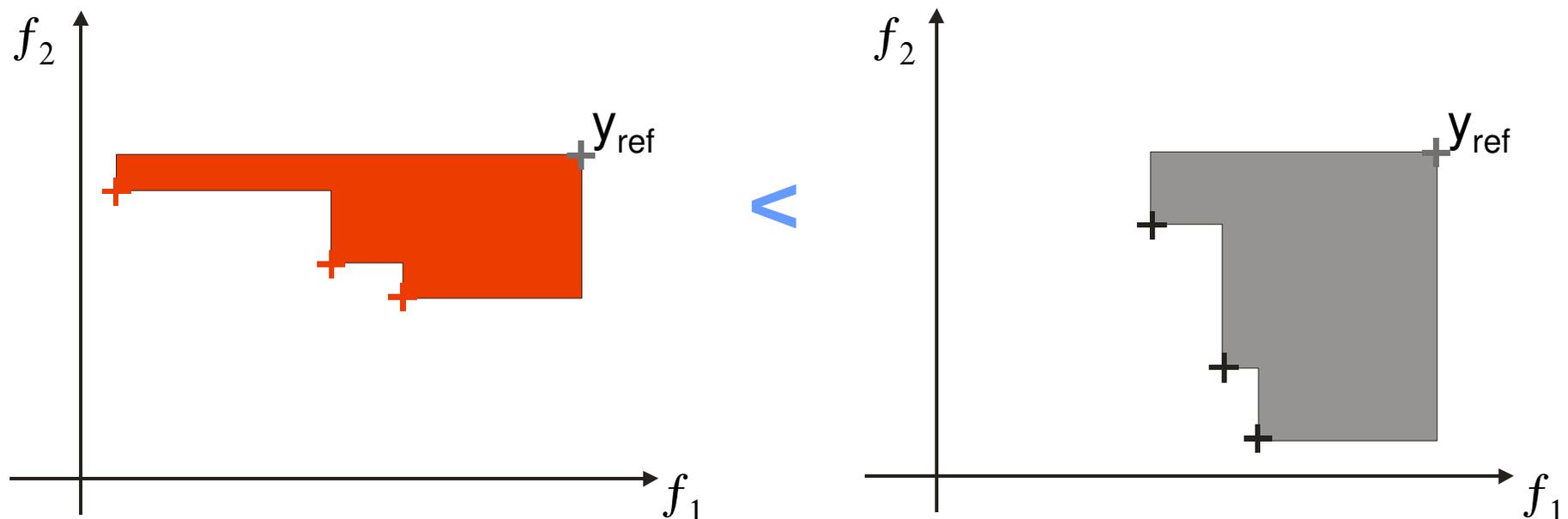
- $S(M) = \max\{S(M)\} \Leftrightarrow M = \text{Pareto-Front}$  (falls PF endlich)
- $A \triangleright B \Leftrightarrow \forall b \in B \exists a \in A: a \preceq b$  und  $A \neq B$   
 $A \triangleright B \Rightarrow S(A) > S(B)$  (einziges unäres Maß)





## Eigenschaften der S-Metrik

- $S(M) = \max\{S(M)\} \Leftrightarrow M = \text{Pareto-Front}$  (falls PF endlich)
- $A \triangleright B \Leftrightarrow \forall b \in B \exists a \in A: a \preceq b$  und  $A \neq B$   
 $A \triangleright B \Rightarrow S(A) > S(B)$  (einziges unäres Maß)
- Bewertung abhängig von Referenzpunkt  $y_{\text{ref}}$





S-Metrik anerkanntes Qualitätsmaß

- hoher S-Metrik-Wert erstrebenswert

Idee: Maximierung während Optimierprozess ansteuern

- integriere S-Metrik-Berechnung in Selektion von EA
- Maximierung von S-Metrik als mono-kriterielle Ersatz-Zielfunktion
- Indikator-basierter EA

Zitzler, Künzli, ETH Zürich 2004

- SMS-EMOA

S Metric Selection Evolutionary Multiobjective Optimization Algorithm

Emmerich, Beume, Naujoks 2005+2006



### SMS-EMOA

Initialisiere Population von  $\mu$  Individuen

**repeat**

wähle 2 Individuen zufällig gleichverteilt

erzeuge 1 Nachkomme durch Rekombination, Mutation

ersetze Individuum mit kleinstem S-Metrik-Beitrag

**until** Abbruchbedingung

Steady-state EA:  $(\mu + 1)$  Selektion

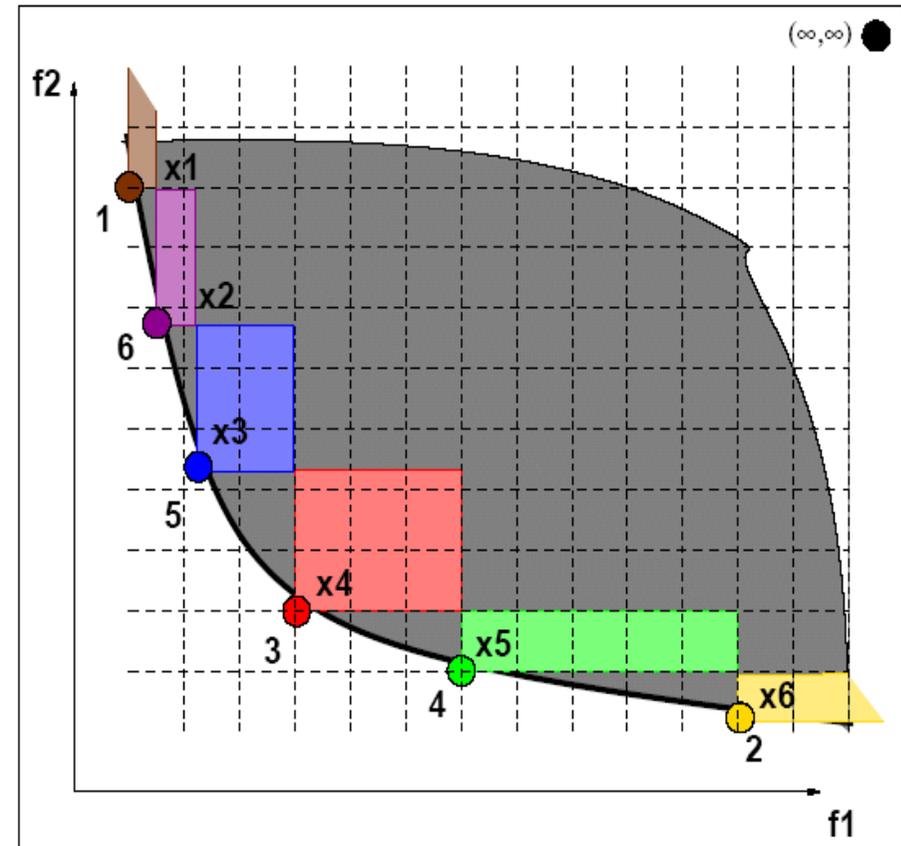
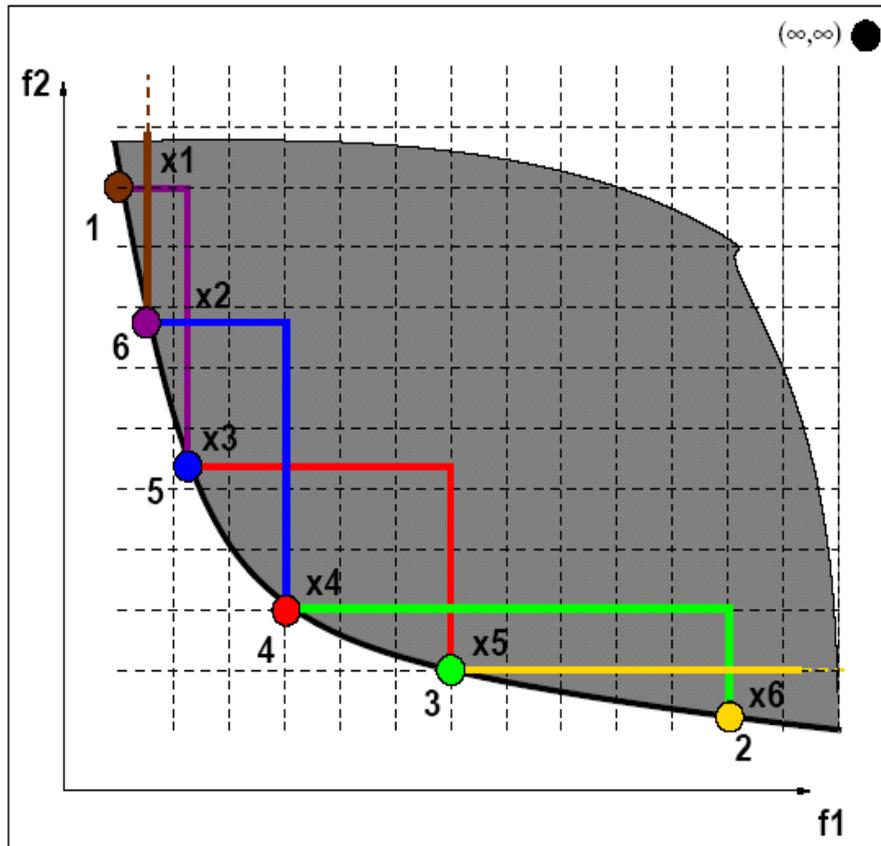
S-Metrik-Wert von Population monoton steigend

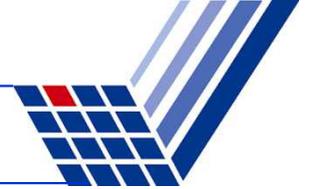
Parameter: Populationsgröße  $\mu$ , Variationsoperatoren



## Vergleich SMS-EMOA und NSGA-II

- Multiplikatives statt additives Kriterium
- S Metrik hängt von Position der Lösungen selbst, nicht (nur) von der Nachbarn ab





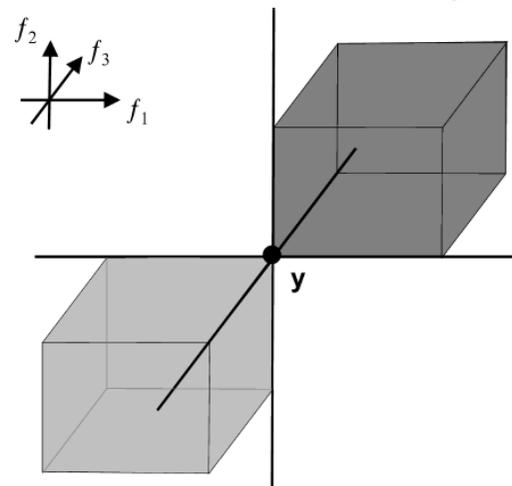
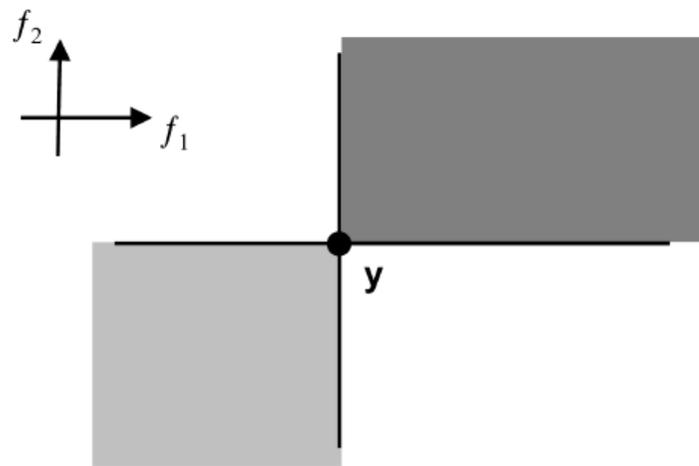
Optimierung hochdimensionaler Funktionen möglich!

relativer S-Metrikwert stabil bei steigender Anzahl Zielfunktionen

NSGA-II, SPEA2: keine Konvergenz zu Pareto-Front

(nur für 2, 3 Zielfunktionen geeignet)

Annäherung an Pareto-Front von SMS-EMOA 100.000-fach besser



Wagner, Beume, Naujoks 2007

Punkt nur mit Anteil von  $1/2^{d-1}$  des Zielfunktionsraums vergleichbar



### Laufzeit-Komplexität von Hypervolumen

- Laufzeit abhängig von Reihenfolge der Punkte, Ziele
- LebMeasure exponentiell in  $k$  (L. While, 2005)
  - bad case:  $O(m^{k-1})$  für 1 Punkt
  - „vermutlich existiert kein exakter polynomieller Algorithmus“
- *allgemein*: Hypervolumen von überlappenden Hyperkuben
- Overmars, Yap (1988):  $O(m^{k/2} * \log_2 m)$  mit plane-sweep
- *hier Spezialfall*:
  - alle Kuben haben Referenzpunkt als „rechte obere Ecke“
  - kein Kubus vollständig in anderen enthalten



Nachteil: hohe Laufzeit

S-Metrik:  $O(\mu \log \mu + \mu^{d/2})$

Beume, Rudolph 2006

SMS-EMOA:  $\mu+1$  S-Metrik-Berechnungen pro Generation

$O(\mu \log \mu + \mu^{d/2+1})$  pro Generation

SMS-EMOA höchst effektiv, an Effizienz wird weiter gearbeitet



### Eigenschaften des SMS-EMOA

- Motivation: kleine, gut verteilte Lösungsmenge
- Steady-state:  $(\mu + 1)$ -Selektion (konstante Populationsgröße)
- Variations-Operatoren: SBX, polynomielle Mutation wie  $\epsilon$ -MOEA (Deb et al., 2003)
- Non-dominated-Sorting aus NSGA-II (Deb et al., 2002)
- entferne Ind. mit kleinstem S-Metrik-Beitrag zum Volumen der schlechtesten Front
- Laufzeit bestimmt von Selektionsoperator
- wenige Parameter: Populationsgröße  $\mu$ , Variations-Operatoren

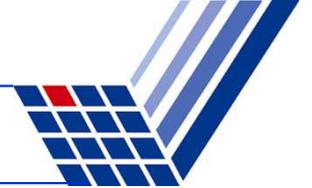


### Eigenschaften der Selektion

- bei  $(\mu+1)$ -Selektion:  $\mu+1$  Mengen vergleichen  
bei  $(\mu+\lambda)$ -Selektion:  $\binom{\mu+\lambda}{\mu}$  Mengen vergleichen
- Greedy-Vorgehen optimal bei Steady-state
- S-Metrik-Wert von Population monoton wachsend
  - Population konvergiert gegen lokales Optimum von S-Metrik  
d.h. Austausch eines Punktes bewirkt keine Verbesserung

Behauptung (Knowles, Corne, 2003):

Punkte einer lokal optimalen Menge „gut verteilt“



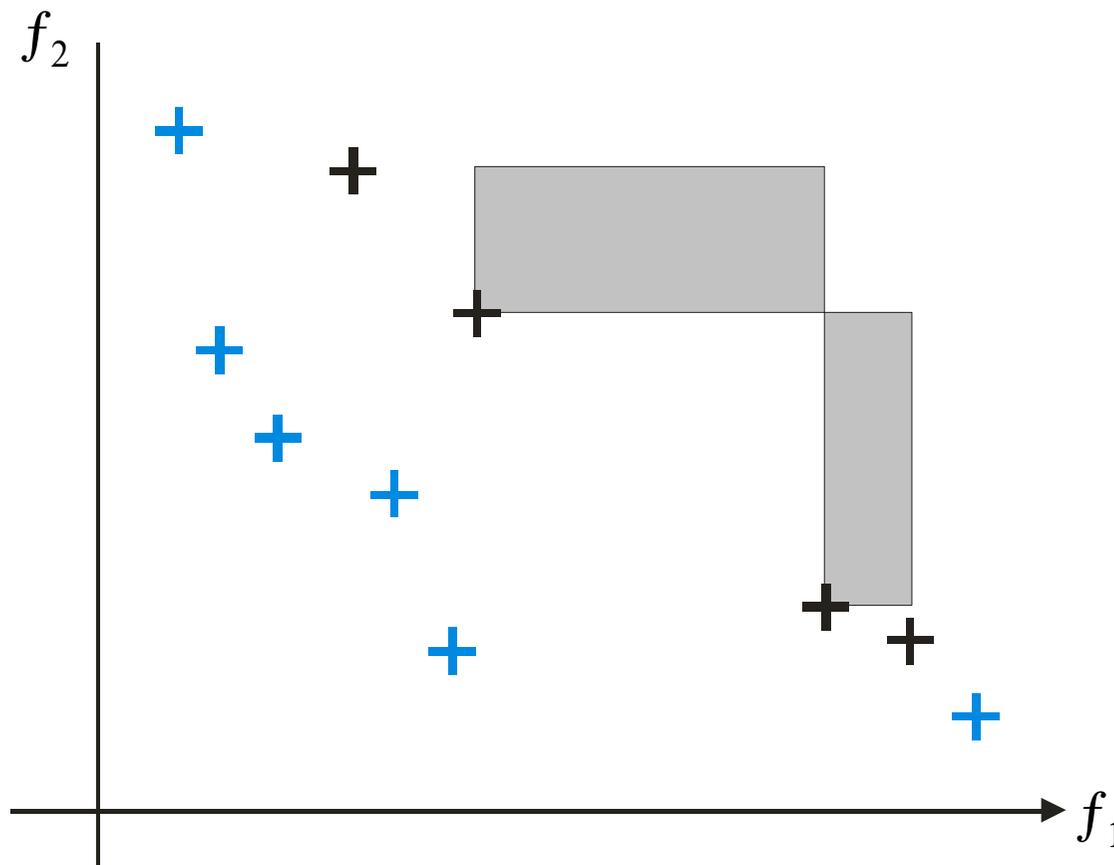
### Empirische Ergebnisse

- Analyse analog zu  $\varepsilon$ -MOEA (Deb et al., 2003)
  - 2 Ziele: ZDT1 - ZDT4, ZDT6, 3 Ziele: DTLZ1 - DTLZ4
  - Bewertung bzgl. S-Metrik, Konvergenz-Metrik
  - Durchschnitt von 5-10 Läufen, Standardabweichung
- ⇒ SMS-EMOA schlägt alle betrachteten EMOA  
(NSGA-II, C-NSGA-II,  $\varepsilon$ -MOEA, SPEA2, PESA)
- Optimierung realer aeronautischer Probleme
- Metamodell-Integration: approximative Funktionswerte



### Modifizierte Selektion

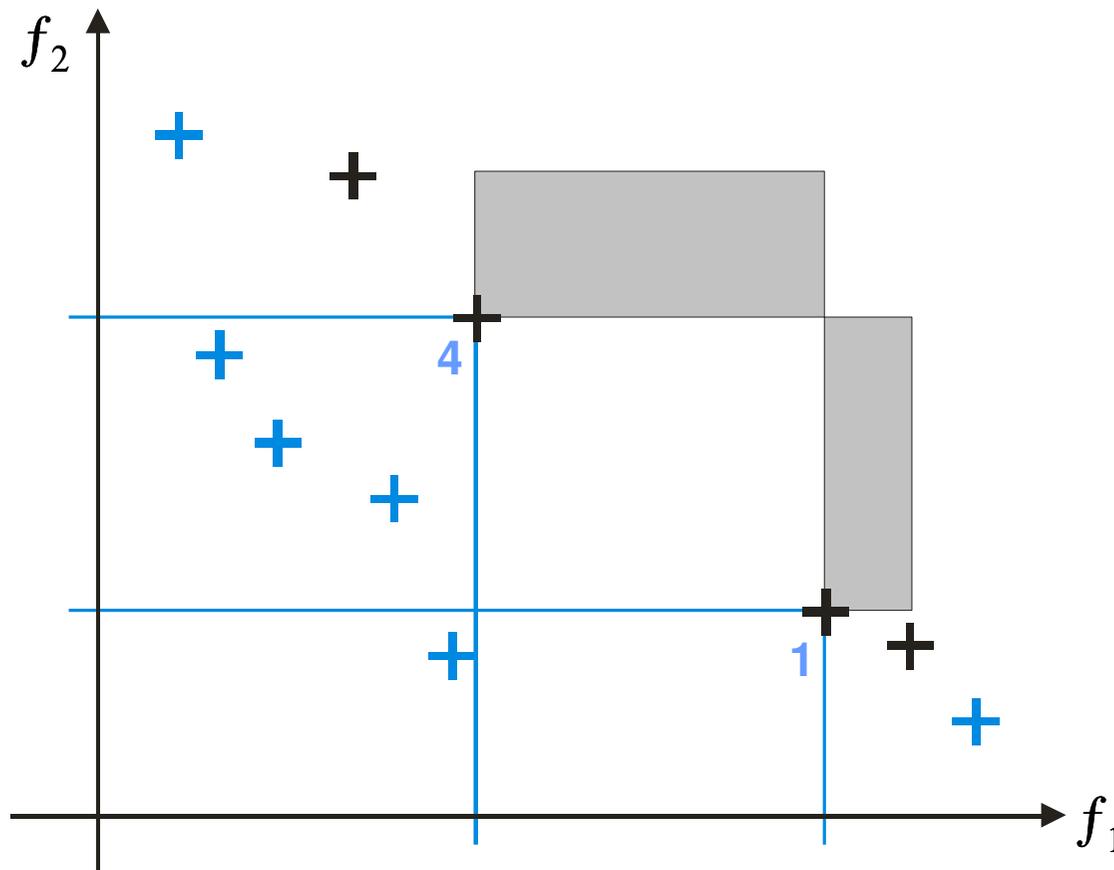
- bisher: gute Verteilung auf allen Fronten angestrebt
- Idee: Selektion abhängig von vorderen Fronten



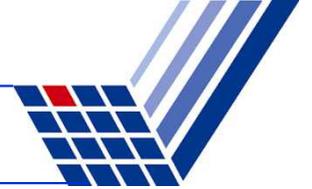


## Modifizierte Selektion

- bisher: gute Verteilung auf allen Fronten angestrebt
- Idee: Selektion abhängig von vorderen Fronten

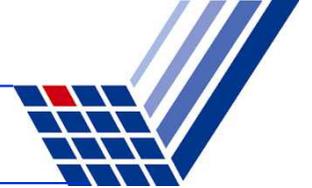


- Maß: Anzahl dominierender Punkte
- anwendbar für dominierte Punkte
- falls #Fronten  $> 1$  berechne  $\Delta_S(v, V_l)$  sonst: #dom. Punkte
- Laufzeit:  $O(\mu * \mu_k)$



### Empirischer Vergleich von Selektionsvarianten

- Qualität von modifizierter Selektion gleich
- wenig schlechter: Hypervolumen & rand. Selektion in  $V_1$
- sehr schlecht: rand. Selektion auf nicht-dominiertes Front
- ⇒ Selektion auf nicht-dominiertes Front am wichtigsten
- ⇒ Laufzeitverbesserung im average case
  
- dynamische Populationsgröße:
  - nur nicht-dominierte Lösungen (schnelle Konvergenz)
  - Ausbreitung auf gesamter Front (Diversität)



### Zusammenfassung

- SMS-EMOA schlägt NSGA-II,  $\epsilon$ -MOEA, SPEA2, PESA in populären Testfunktionen, realen Anwendungen
- Hypervolumenberechnung zeitaufwändig: exponentiell  
bzgl. #Ziele  $k$
- SMS-EMOA geeignet für
  - zwei-kriterielle Zielfunktionen
  - rechenintensive Funktionsauswertungen
- effiziente Selektion unter dominierten Punkten nach  
Anzahl dominierender Punkte



### Weitergehende Fragen

- Metaheuristiken-Entwurf
  - Indikator-basierte Selektion: weitere Indikatoren?
  - S-Metrik-Selektion in anderen Algorithmen (z.B. PSO)?
- Empirische Analyse von EMOAs
  - günstige Parametrisierung des SMS-EMOA?
  - Anzahl dominierender Lösungen statt Non-dom. Sorting?
- Theoretische Analyse
  - Eigenschaften lokaler Optima der S-Metrik?
  - Konvergenz und Optimierzeit von SMS-EMOA?



### SMS-EMOA

- Steigerung der Effizienz: Laufzeitverbesserung durch effiziente Updates zwischen Generationen
- Studie der Verteilung der Ergebnismenge
- Implementierung in versch. Programmiersprachen, Frameworks



### Mehrzieloptimierung allgemein

- Richtlinien zur Anwendung von EA bei versch. Optimierszenarien
- Zusammenhang Pareto-Front  $\leftrightarrow$  Pareto-Menge  
z.B. finde Lösung mit gleichen Funktionswerten und einfacheren Objektparametern

Rudolph, Naujoks, Preuß 2006+2007

- Parameterstudie über versch. EMOA
- Implementierung in versch. Programmiersprachen, Frameworks



### Literatur

#### ■ Bücher:

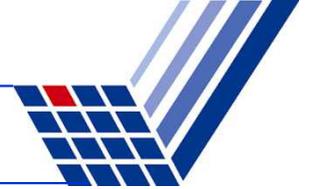
- Deb: **Multi-Objective Optimization using Evolutionary Algorithms**, John Wiley & Sons, Chichester, UK, 2001
- Coello Coello, van Veldhuizen, Lamont: **Evolutionary Algorithms for Solving Multi-Objective Problems**, Kluwer Academic Publishers, New York, 2002

#### ■ Übersichtsartikel

- Fonseca u. Fleming: *Evol. Comp. Journal (MIT)*, 1995

#### ■ Originalarbeiten zu vorgestellten Algorithmen

- Email an [boris.naujoks@uni-dortmund.de](mailto:boris.naujoks@uni-dortmund.de)



# Multikriterielle Optimierung mit evolutionären Algorithmen

## Zusammenfassung

1. Beispiel
2. Historie
3. Grundlagen
4. Einfache, aggregierende Ansätze
5. Standardverfahren NSGA-II,  $\epsilon$ -MOEA
6. Vergleich von Ergebnissen: Metriken
7. SMS-EMOA
8. Ausblick, offene Fragen etc.